HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

The differences between bacterial communities in white clover (*Trifolium repens*) in rural and urban areas

University of Helsinki
Faculty of Biological and Environmental Sciences
Environmental Change and Global Sustainability - ECGS
Master's thesis
Huhtikuu 2021
Iida Mäkelä

UNIVERSITY OF HELSINKI

Tiedekunta – Fakultet – Faculty	Koulutusohjelma – Utbildningsprogram – Degree Programme	
Bio- ja ympäristötieteellinen tiedekunta	Environmental Change and Global Sustainability (ECGS)	
Tekijä – Författare – Author		
Iida Mäkelä		
Työn nimi – Arbetets titel – Title		
The differences between bacterial communities in white clover (Trifolium repens) in rural and urban areas		
Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track		
Environmental Change		
Työn laji – Arbetets art – Level	Aika – Datum – Month and year	Sivumäärä – Sidoantal – Number of pages
Master's thesis	Huhtikuu 2021	27

Tiivistelmä – Referat – Abstract

Microbial diversity can be found everywhere around us. The diversity is however declining globally and the diversity loss is most visible in highly urbanized areas. The lack of microbial biodiversity has been linked to increased risk of certain immune mediated diseases most prevalent within urban population. Understanding how diversity differs between urban and rural areas can help us to figure out mechanisms behind biodiversity loss and higher frequency of immune-mediated diseases and develop prevention methods for the latter.
The aim of the thesis is to study how bacterial communities differ between urban and rural areas using indicator species as proxy. The aim is also to find out if the results support the biodiversity hypothesis.

The results of the thesis found out significant differences in diversity indexes between bacterial communities in urban and rural areas, which supports the biodiversity hypothesis. The study also found differences in Proteobacteria diversity indexes, which have been linked to some immune mediated diseases in previous studies.

Avainsanat – Nyckelord – Keywords

Trifolium repens, urbanization, biodiversity hypothesis, microbe, diversity, hygiene hypothesis

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors

Laura Soininen, Aki Sinkkonen, Marja Roslund

Säilytyspaikka – Förvaringsställe – Where deposited

Viikin tiedekirjaston E-thesis tietokanta

Muita tietoja – Övriga uppgifter – Additional information

Table of contents

1 Introduction

Microbes can be found everywhere. They colonize nearly all surfaces on earth and live as commensal communities inside other organisms, like plants and animals. The area of science that studies microbes is called microbiology. It was discovered after the invention of a modern microscope, which enabled scientists to observe biodiversity that would otherwise be mostly invisible. Another important achievement for microbiology is the modern DNA sequencing techniques, which enable the detailed studying of genes from different organisms. This allows scientists to identify and study the microbial diversity and their taxonomy.

Microbes have the ability to survive in extreme conditions, where other organisms would not be able to survive. Extreme environmental conditions may include radiation, fluctuating temperatures or pH levels, or all of the above (Pepper et al., 2011). Microbes are typically the pioneer colonizers in the environment, which means that they will inhabit new areas first. If there is free living space available, it is most probably already inhabited by microbes. The term microbe covers different types of microorganisms that are classified as microbes by their microscopic size. Microbes include bacteria, archaea, fungi, protozoa, algae, where bacteria is the biggest group calculated by its total biomass (Bar-On, Philips & Milo, 2018). Viruses can also be considered microbes based on their size, but it is important to note that they are not in fact microorganisms, since they are not classified as living. (Pepper et al., 2011).

Another important group of organisms is plants, which are closely interlinked with microbes. Plants are the biggest kingdom of biosphere and together with microbes; they form approximately 96% of the total biomass on Earth's biosphere (Bar-On, Philips & Milo, 2018). The ecological functions provided by microbes and plants are necessary for all other life forms on earth. They transform the energy from the sun into an organic form, which other organisms are able to consume. They are also responsible for cycling the nutrients between organic and inorganic worlds (Pepper et al., 2011). Together plants and microbes create an environment where all other species, like mammals, can flourish.

The overall biodiversity has been declining rapidly during 19th century due increased human population and industrialization. This trend is most visible in highly urbanized areas, where human activity is the most intense (Piano et al., 2020; McK). This phenomenon - the biodiversity hypothesis, was introduced by Von Hertzen et al. in 2011. It points out that biodiversity is less diverse in urban than rural areas. This decline in diversity concerns also microbes (Pepper et al., 2011). Decline in microbial diversity can lead to ecological problems, as microbes are responsible for a variety of important functions within ecosystems (Tortora et al. 2004; Pepper et al., 2011). It has also been noted that environmental microbial diversity (especially bacterial diversity) in densely populated areas is linked to human health (Macpherson and Harris, 2004; Round and Mazmanian, 2009; Roslund et al., 2020). Urbanization and biodiversity loss create an environment for humans, which cannot provide sufficient stimuli for our immune system. This lack of stimuli from our surroundings and the increased incidence of autoimmune diseases was linked to high hygiene levels in 1989 by an epidemiologist Dr. David Strachan, who formed the hygiene hypothesis. It has also been observed that people living in an urban environment tend to spend less and less time in the nature. This means that exposure to rich microbial diversity has declined as urbanization continues (Parajuli et al., 2018; Hui et al. 2019).

There is plenty of evidence that environment and especially bacteria have a significant role in the development and the maintenance of a healthy human immune system (Macpherson and Harris, 2004; Round and Mazmanian, 2009; Roslund et al., 2020; Von Hertzen & haahtela, 2006; Von Hertzen & Haahtela, 2006). Diverse plant community around one's home seems to alter gut microbiota towards a more balanced state (Parajuli et al., 2020). This is likely because a diverse plant community can provide a good habitat for a rich microbial community. People who spend a lot of time in the nature are in close contact with these microbe rich biofilms and the diversity of our own microbiome can be enriched by being in physical touch with microbe rich environment, such as soil (Parajuli et al., 2020; Grönroos et al. 2019; Roslund et al. 2020). Studies point out that rich biodiversity around us can provide beneficial exposure to microbial diversity that can lower the risk of some immune-mediated disease, including diseases such as type 1

diabetes, atopy and asthma (Von Hertzen et al. 2011;Strachan, 1989; Hanski et al.2012; Strachan, 2000; Ruokolainen et al. 2016; Kondrashova 2012; Kostic et al. 2015; Von Hertzen & Haahtela, 2006). They can give a comprehensive idea how we should design our surroundings, or in other way try to add more biodiversity in to our lives inside the city borders.

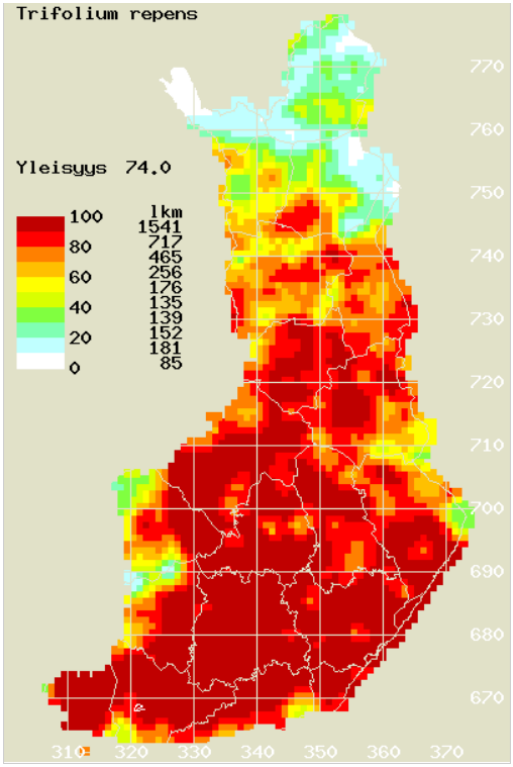
Immune system's relationship with microbial stimuli from our surroundings can be explained as "The Old Friend Hypothesis", which supports the hygiene hypothesis described in 1989 (Strachan, 1989The Old Friend Hypothesis suggests that humans have coevolved with abundant but harmless saprophytic bacteria (e.g. mycobacteria and lactobacilli) that break down dead plant matter from the environment. Water and food that humans consumed included these bacteria and our immune system needed to tolerate their presence in and on our bodies. This stimuli coming from harmless environmental microbes suppress other inflammatory responses as our body recognizes this type of stimulation as harmless (Rook, 2009), has been supported by many studies and clinical trials (Manlius. et al., 2002; Wilson et al., 2005; Ricklin-Gutzwiller et al., 2007; Zaccone et al., 2003). Immunological responses trying to fight against these common bacteria would have most likely resulted in severe tissue damage following other problems without modern medicine (Rook, 2009). Mitigated contact with other common bacteria like Proteobakteria has also been linked to immune system-related diseases (Hanski et al., 2012). Studying how the bacterial diversity has changed in our environment during urbanization is important for understanding fundamental environmental and health-related connections.

In order to understand diversity loss within microbes it is necessary to determine "who" are present in the samples. This is done by modern sequencing techniques, which target specific area of genetic code that is only present in kingdom of interest. For bacteria, this target gene is the 16S, which has stayed the same throughout the bacteria kingdom. Amplifying genes including this segment target specifically bacteria. (Knight, 2018; Turner et al., 2013). This genomic data is then compared to existing phylogenetic database, which can give identification for the bacteria in question.

Most commonly used calculating methods for species diversity in a sample are Shannon diversity index and Simpson's concentration index. These formulas indices indicate how many different species are observed in the sample, as well as its relation the overall abundance of species in the community. The basic principle of Shannon Index is that individuals of one species is divided by the total number of individuals found in the sample. In addition, the Simpson index that gives more weight to the most abundant species. Basically, Simpson's index determines how probable it is that two individuals that are randomly chosen from the community, are same species. (Lande, 1996). Calculating species diversity has a major role in many ecological studies. Determining species abundance and richness at various taxonomic levels and comparing that information with distributional information gives scientific information how differences in the environment can affect microbiological abundance. This information can also be approached by deconstructing taxonomic data into smaller groups. This approach, known as deconstructed species diversity, examines taxonomic data in smaller groups to detect community patterns that might be undetectable in a whole community scale (Marquet et al., 2004). Differences on specific levels can tell detailed information about the change in the environment, since the main cause of diversity loss in Phylum and Species (which is referred as OTU or Operational taxonomic unit in bacteria) level can be different. This information can help scientist to determine and understand the underlying issues behind biodiversity loss, which starts at the microscopic level and extrapolates to all life on earth.

In science, many phenomena are studied and described by using model species for practical and financial reasons, as it is impossible to take into account all species in the area at the same time. If scientists want to observe a small-scale transition, it is reasonable to choose an indicator species to study, as some species are sensitive for even a small change in the environment. On the other hand, if a scientist wants to study more general phenomena happening in a bigger scale, it is useful to choose a species that is more or less average representation of an abundant species. When comparing diversity between urban and rural areas, it is important to choose a species that can thrive in both habitats. White clover *Trifolium repens* is a good model species in studies along the urban and rural-axis, as white clover can tolerate many different conditions and it is widely used

in commercial lawns used in both city landscaping and on private yards on countryside, as well as pastures.



White clover (Trifolium repens) is a legume that is widely used in lawns and pastures, but it also grows naturally on roadsides and wastelands worldwide. It can tolerate many climate conditions and nutrient poor soils by fixating nitrogen from the atmosphere (Frame & Newbould., 1986). It is about 10-30 cm tall, creeping perennial that flowers during June-August. It can be found in many different locations due its tolerance of many different conditions. (Luontoportti, 2021)

Figure.1 Prevalence of Triflium repens in Finland (Source: University of Helsinki, 2021)

Plants and microbes have a close relationship with each other. Microbial communities depend on plant part,as well as on environment where the plant is living. The root system (rhizosphere) and aboveground part of the plant (phyllosphere) work as different environments for microbes. The aboveground part of the plant is in direct contact with organisms living on ground and as well as the microbes on the plant. Humans or other animals that are in direct contact with plant leaves by touching or eating them are also in contact with the microbial community living on the leaves. Understanding the differences in bacterial diversity on *Trifolium repens* leaves between urban and rural areas can give us understanding how different exposures to biodiversity people have depending if they live in a city or countryside.

2 Theoretical Background

In order to study how *Trifolium repens* associated bacterial communities differ between urban and rural areas, it is important to understand more about plant microbiome. Plants host microbial communities throughout their structures. Diverse microbiota has many benefits and it has been studied they promote plant health, reproduction and survival (Turner et al., 2013). All plant species have a unique blend of beneficial, commensal and pathogenic bacteria. These communities are dependent on other species and their abundance, but also on the environmental factors locally (Berg and Smalla, 2009). The rhizosphere is (mainly) soil derived and the bacterial diversity is related to environmental factors as well as exudates from plant roots. In the rhizosphere, most dominant groups of bacteria are Proteobakteria, Actinobacteria, Firmicutes, Bacteroidetes, Planctomycetes, Verrucomicrobia and Acidobacteria (Turner et al., 2013). From those Actinobacteria, Acidobacteria, Proteobakteria, Bacterioidetes, and Firmicutes are the most prevalent groups in the soil outside rhizosphere (Fierer et al., 2007). Phyllosphere is on the other hand, a different living environment for microbes compared to the rhizome: it barely has any nutrients and it is subjected to fluctuating temperatures, radiation and moisture levels. Microbes can colonize it by up to 10^7 microbes per cm^2 (Vorholt, 2012; Turner et al., 2013; Lindow & Brandl, 2003). Urbanization creates a living environment for plants and associated microbes that is often characterized by only few species, high intensity trampling, heat island effect and pollutants (Parris, 2016; Roslund et al., 2018)

The aboveground part of the plant (endosphere and phyllosphere) is the main focus in this study, as it is the part of the plant that is in contact with humans. The phyllosphere is mainly dominated by Proteobakteria, Bacteroidetes, Actinobacteria, Firmicutes, but the dynamic environment it provides for the microbes can cause alteration for microbial composition depending on the external conditions (Turner et al., 2013; Vorholt, 2012). The same phyla of bacteria on leaves are present in the soil (Fierer et al., 2007), although there are many differences on lower taxonomic levels and the overall diversity is reduced (Vorholt, 2012). The

endosphere inhabited by endophytic bacteria are considered to be a sub-population of the rhizosphere microbes, although they have distinctive characteristics apart from rhizosperic bacteria (Compant et al., 2010). Certain bacterial phyla that can be found on plant leaves have also been linked to studies that focus on the bacterial communities and health in humans (Rook, 2009; Hanski et al., 2012). Interesting phyla are Proteobakteria as well as Firmicutes and Bacterioides: Firmucutes;Bacterioidetes ratio was altered after test subjects touched soil enriched with microbial inoculant (Hui et al, 2019).

2.1 Aims of the study

In this study, the main goal is to find out how bacterial communities that live in and on white clover (*Trifolium repens*) differ from each other based on the location of the plant. The aim is to find out how bacterial communities in urban and rural areas differ from another and create more understanding and knowledge of how urban and rural environments differ on a microscopic level. The hypothesis was that biodiversity of phyllosphere was higher in rural than urban locations.

Although there are multiple studies about urban biodiversity, this is one of the first studies that observe this phenomena using an indicator species. This study was conducted using leaves and stems of white clover (*Trifolium repens*). It was chosen as the indicator species for this study, because it can be easily found in urban and rural landscape and the study can be possibly replicated in many other locations. I focus on differences between bacterial communities living on a white clover between urban and rural environment to gain more information of the biodiversity hypothesis (Von Hetzen et al, 2011). Biodiversity loss of microbiota of an indictor specie can further support the Hygiene hypothesis (Strachan., 1989) and link biodiversity loss at a bacterial level to The Old Friend Hypothesis (Rook.,2009), as well as the prevalence of auto-immune diseases within urban population (Hanski et al.2012; Strachan, 2000; Ruokolainen et al. 2016; Kondrashova 2012; Kostic et al. 2015).

3 Materials and methods

3.1 Study area and sampling

During this case study, the main goal is to find out if microbial communities in phyllosphere and endosphere are different between urban and rural areas. The design and protocol of the study follows other studies by ADELE research group (e.g. Grönroos et al., 2019; Parajuli et al., 2020). Microbial analysis is done using 16S sequencing method and the data tested with statistical tests commonly used in ecological studies (Knight, 2018; Turner et al., 2013). Species diversity and diversity indices are then compared between these two variables (urban/rural) to determine if there are differences on bacterial diversity between these environments.

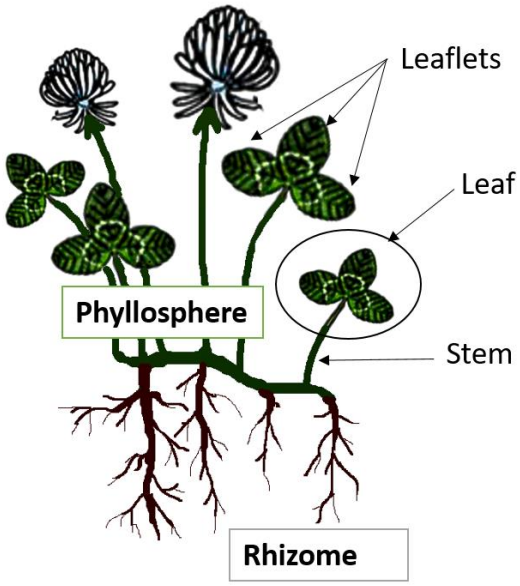


Figure 2. Anatomy of white cover (Trifolium repens)

Trifolium repens samples were collected from urban area (Lahti) and from countryside (municipalities in Päijät-Häme and regions close proximity) during summer of 2015. More specific areas within a map can be found listed in the references (Appendix 1). The sampling site in each address covered 10 acres, and sampling points were chosen with at least six meters between sampling spots.

Nine samples were collected from each yard i.e. sampling site. Each sampling site had three randomly chosen areas of 2 meter radius. Within each area three clover samples were collected. Each sample was from different spot to avoid covering the same plant (connected in rhizome) multiple times. If *Trifolium repens*

could not be located from the original randomly chosen sampling spot, a new sampling spot was chosen at random. Samples were taken from phyllosphere of *Trifolium repens*, the above-ground-part of the plant, with scissors and tweezers that were sterilized with 70% ethanol. This reduces the risk of contamination when handling the sample. Collected samples were sealed in a zipper storage bags and the bags were immediately transferred into a cold storage. The total number of sample bags collected per address was thus nine.

3.2 Sample preparation and laboratory analysis

Trifolium repens samples were stored in a deepfreezer (-80ºC) in the same zipper bag the samples were originally collected in, after which they were retrieved for sampling. Stored sample bags contained 3 leaves with stem and the final sample included one leaflet and piece of stem of each of these plants (Figure 2). Samples were transferred into 1,5ml collecting tubes and weighed before transferring samples into DNA extracting tubes provided by DNeasy PowerSoil Pro Kit. During the laboratory analysis the samples were processed in random order to mitigate the risk of batch bias for the final data.

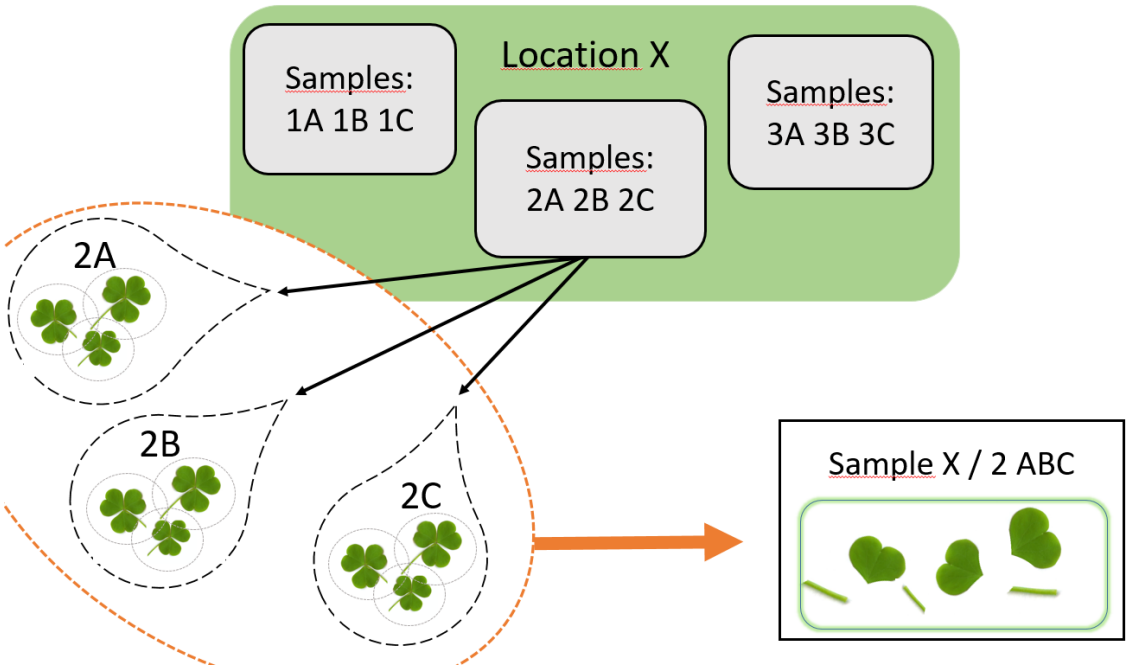


Figure 3. Illustration of sampling in situ and how the samples for DNA extraction were composed.

The bacterial DNA from *Trifolium repens* was extracted using commercial DNeasy PowerSoil Pro Kit by QIAGEN. This method has been used in similar research publications from ADELE- research group (Grönroos et al., 2019). The protocol provided with PowerSoil kit was used as manufacturer's standard instructs. Before the bacterial DNA could be extracted it was necessary to break down the plant material to include also bacteria from endosphere. Plant material was broken down in PoweBead tubes, which were included in the kit. Individual tubes contained approximately 150 µg of plant matter, which was measured before applying it into PowerBead tube. After extracting bacterial DNA from the plant material,DNA concentration was measured using Quant-iT™ PicoGreen® dsDNA reagent kit manufactured by Thermo scientific (MA, USA).

The amplification of bacterial DNA was done targeting 16S region V4 of ribosomal RNA (rRNA). PCR (polymerase chain reaction) amplification used 505F and 806R primers (Caporaso et al., 2012). Original samples were triplicated before PCR. The reaction mixture for PCR was built from 1 µl each of 10 mmol/L dNTPs (Thermo scientific, MA, USA), 5 µl forward primer 505F (10 µmol/L; 5′–

GTGCCAGCMGCCGCGGTAA-3′) and 5 µl reverse primer 806R (10 µmol/L; 5′–GGACTACHVGGGTWTCTAAT-3′), 0.5 µl 2 U/µl Phusion Green Hot Start II High-Fidelity DNA polymerase (Thermo scientific, MA, USA), 10 µl 5× Green HF PCR buffer (F-537), 5 µl template DNA and added 23.5 µl of sterile water to the mixture for total volume of 50µl for reaction.

PCR was conducted in a thermocycler (MJ Research, MA, USA) using following steps: first denaturation at 98°C for 5 min, following 30 cycles with denaturation at 94°C for 1 min, 10 s at 50°C , 1 min at 72°C, and at the final step 72°C for 10 min. During these cycles, DNA is replicated and the concentration doubles during each cycle. During heating phase, DNA chains separate into strands, where primers from a mixture attach themselves to form two identical chains. The successful working of the PCR was confirmed using a bacterial community that had been previously extracted as positive and autoclaved water as negative controls in each batch of 12x8 microwell plate. The resulting PCR products were purified with Agencourt AMPure to avoid transferring PCR primers onwards. Lastly, trifold samples were combined into one sample, which was sent to Finnish Institute for Molecular Medicine (www.fimm.fi) for processing the sequences with Illumina MiSeq platform (Ravi et al., 2018).

3.3 Bioinformatics

The sequence data from Meilahti was processed with Mothur program. Mothur was operated using CSC's Chipster version 3.15 operating program (Kallio et al., 2011), which is Mothur based and open source visual operating system for sequence data. The operational functions are therefore the same as in Mothur (version 1.43.0). Sequence data goes through the standard operating procedure (SOP), for 16S rRNA sequence (Schloss et al., 2009; Schloss et al.,2011; D'Amore et al., 2016; Pollock et al., 2018) and followed the pipelined that was suggested by Schloss et al., 2011 and Kozich et al., 2013)

First step was to pair raw sequences that are divided at this point into forward and reverse fastq files and then align them into contigs. After pairing fastq files they were trimmed and mutated sequences were removed alonside ambiquous bases or >8 base pair (bp) long homopolymers. Trimmed and screened sequences were compared into Mothur version of SILVA v.132 bacterial reference sequences (Quast et al., 2013) and unaligned sequences that couldn't be found from the database removed. This database identifies and give names to bacterial taxons based on the sequence data that is compared to it. The number of unique sequences was checked from the data and identical sequences with >99% similarity were preclustered together. This procedure was done to mitigate possible risk of sequencing error during the process (Huse, Welch, Morrison, & Sogin, 2010). Sequence data was then screened for chimeric sequences, which were detected with UCHIME (Edgar et al. 2011). UCHIME uses most abundant species as a reference when it detects chimeras, so they can be removed from the raw sequence data.

The remaining sequences were clustered into OTUs (Operational taxonomic units) using 97% sequence similarity as a reference value. At this point all the sequences belonging to bacteria were saved and other kingdoms including chloroplast, mitochondria, Archaea, Eukaryota and unknown sequences were removed from the dataset. OTUs that were rare (<10 sequences) were removed, to mitigate false diversity created by mispaired sequences during PCR amplification (Oliver et al. 2015).This final output from Mothur provides taxonomic data from clover samples that can be statistically tested. (Schloss et al., 2011; Knight et al., 2018). The files containing taxonomical information were manipulated with Excel 2016 (Microsoft Corporation. (2020). *Microsoft Excel*) and JMP Pro 14 (SAS Institute Inc., Cary, NC, 1989-2020.) programs, to format them into correct form for further statistical analysis. Samples were then subsampled to 2549 sequences, which was the smallest amount sequences found in a sample belonging to the dataset. These values were then checked with Good's coverage index (0.96+/-0.02), which indicates how well the subsampling limit covers the total population in further analysis.

3.4 Statistical methods

Taxonomic levels (phylum, class, order, family, genus and OTU), Shannon diversity index and Simpson's evenness were compared between rural and urban areas. Analysis was conducted using R (version) and furthermore using package vegan (version) for diversity analysis and community ordination commonly used in ecology (Oksanen et al., 2017).

Normality was calculated for samples using Shapiro-Wilk test (Shapiro and Wilk, 1965), in order to determine if p-values are calculated with T-test, which follows normal distribution curve, or non-parametric Wilcoxon test for samples which doesn't follow normal distribution (Wilcoxon, 1992). This is important for reliability (Ghasemi and Zahedias, 2012). Afterwards the p-values are corrected with false discovery rate "FDR" method Benjamini and Hochberg into Q-values (Benjamini and Hochberg, 1995). The FDR-corrected Q-value indicates the difference in diversity between urban and rural areas in R. This computing for Shannon Diversity index and Simpson's index was done with *diversity* function from the vegan package (versio). Formulas for the diversity indexes are:

Shannon Index $= -\sum_{is=l}^{s} p_i \ln p_i$

Simpson Index $= \frac{1}{\sum_{i=l}^{s} p_i^2}$

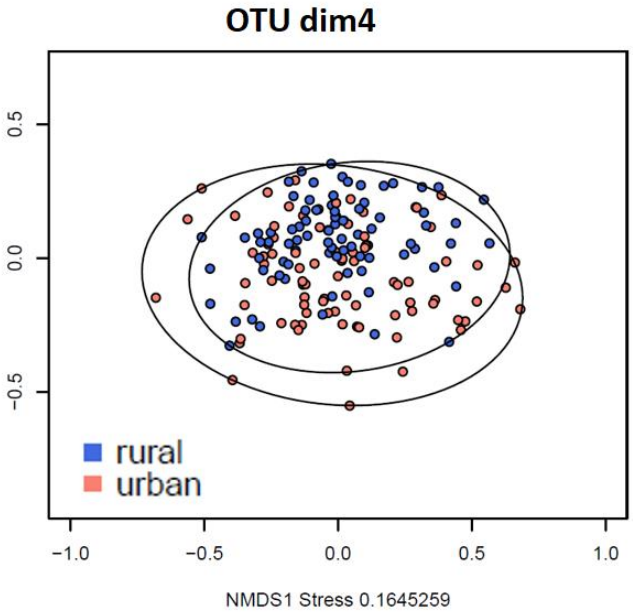
(where *p* is frequency of species (*i*) in a community, *S* is number of species).

The bacterial composition was examined with Permutational Multivariate Analysis of Variance, which is in short PERMANOVA (Anderson, 2001). PERMANOVA was run in R using 999 permutations, which are re-arrangements of gene in this case. The results was then visualized with Non-Metric Multidimensional Scaling (NMDS) using metaMDS (Nonmetric Multidimensional Scalig With Stable Solution From Random Starts, Axis Scaling And Species Scores), which helps to summarize the data and make it easier to see the possible differences between microbial communities. In R *vegan* package Multivariate Analysis of Variance is called ADONIS, which parts sums of squares of a dataset that is multivariate, computing the same multivariate analysis of variance as MANOVA (Anderson, 2001; Anderson, 2005.) PERMANOVA and NMDS coordination analysis was based on Bray-Curtis distance (Bray and Curtis, 1957)., which was calculated for relative abundance data. Analysis was also done separately for phylum Proteobakteria.

4 Results

4.1 Bacterial community characterization

The overall bacterial community composition between urban and rural areas had no significant differences at any taxonomic levels (Phylum, Class, Order, Family, Genus, OTU), as it was examined using permanova and visualized through NMDS. The same protocol was also done for Proteobakteria, and the results were similar with the whole community diveristy.



Although the overall bacterial community characterization were similar with both variables in NMDS plots, there are significant differences in diversity indexed between urban and rural areas, which will be discussed on the next chapter..

Figure 4. NMDS scale of OTU level diversity between urban and rural areas

4.2 Diversity indexes between urban and rural areas

Although bacterial community composition was similar based on permanova result the NMDS plot, the overall diversity showed significant differences in diversity indices calculated in this study. Rural areas had more diverse bacterial community based on Simpson's evenness and Shannon diversity index (Figure 6). This same trend is visible also in bacterial phylum Proteobacteria, which was more diverse in rural areas (Figure 7).

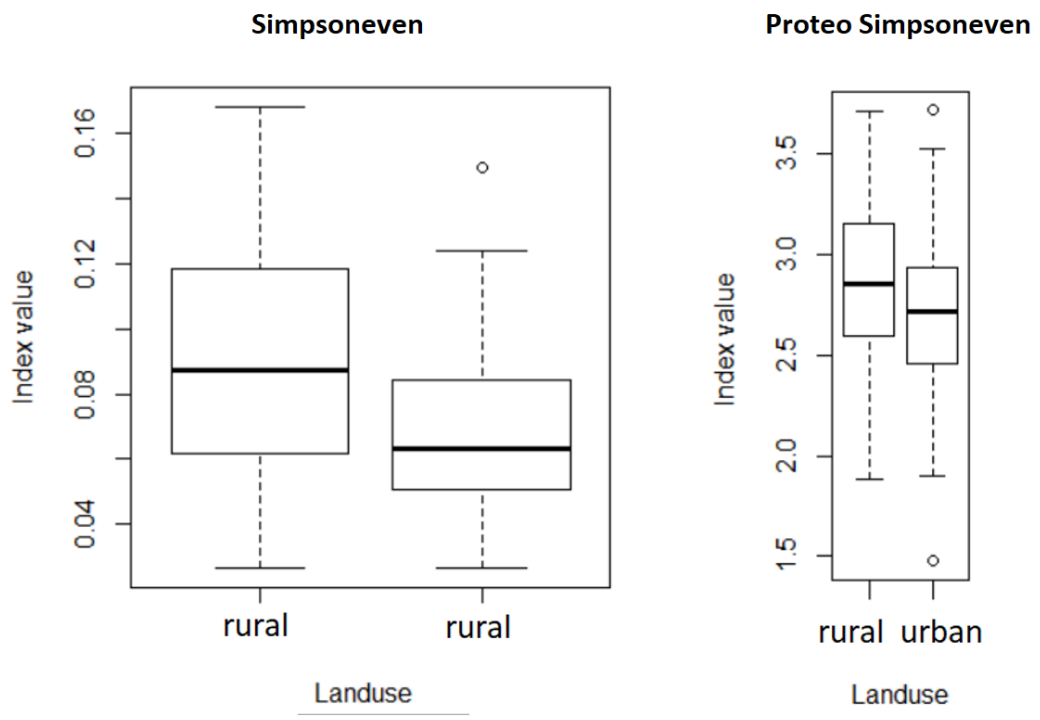


Figure 5. Simpson's diversity index results from bacterial communities from urban and rural areas (Q=0.0008). On the right same index on phylum Proteobakteria (Q=0.017)

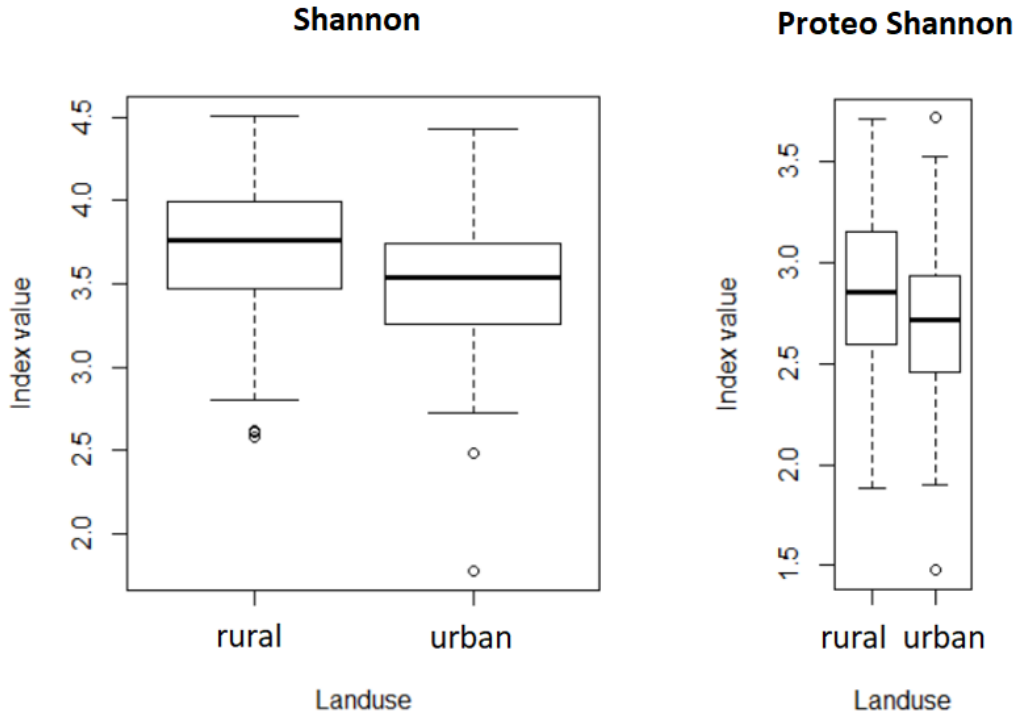


Figure 6. Shannon diversity index results from bacterial communities from urban and rural areas (Q=0.0012). On the right same index on phylum Proteobakteria (Q=0.014)

5 Discussion

The results from this study supported the hypothesis that urban environment reduces biodiversity and further support the biodiversity hypothesis (Von Hertzen et al. 2011). The original hypothesis at the beginning was to determine if microbial communities differ from each other between rural and urban areas. As this is one of the first studies specifically on aboveground plant microbiome between rural and urban areas, there is no other studies to compare the results, so possible extrapolation of the findings is still unknown. The findings in this study however follows other studies and their results (McKinney et al., 2002; Piano et al.,2020; Rook et al., 2009; Von Hertzen et lk., 2011). The lack of microbial diversity in urban areas on *Trifolium repens* therefore supports the biodiversity hypothesis and other studies on microbial loss in an urban environment (Von Hertzen et al. 2011; Grönroos et al., 2019; Hanski et al., 2012; Hui et al., 2019; Lehtimäki et al., 2017; Lehtimäki et al., 2018; Ley et al., 2006; Roslund et al., 2020;Sinkkonen et al., 2018; Strachan, 1989).

This study tested the use of a single plant species as a proxy of biodiversity in the living environment of elderly people studied earlier in Nature-Based Solutions research group (Parajuli et al. 2018, Hui et al. 2019, Parajuli et al. 2020). The results demonstrated that the microbiota on the indicator species was poorer in densely built urban environments, compared the surrounding countryside. While the earlier studies found an association between the biodiversity carried indoors and surrounding land cover types, the current study enlarges the findings to biodiversity in identical habitats at the yards.

The study was conducted in Southern Finland in a city called Lahti. It is 8th biggest city in Finland in 2019 (Kuntaliitto) with approximately 120 000 inhabitants. Lahti identifies as a “green city” and focuses on conserving green areas inside the city. Lahti is built next to lake Vesijärvi lake and the shoreline is filled with parks that local people often visit. In Finland cities are relatively small compared to other European countries, so the close proximity of nature contributes to biodiversity within city borders. Despite this, the current study indicates that even the small and “green” cities have a lower bacterial biodiversity compared to the surrounding

countryside. This finding supports the MacKinney's study (2002) that pointed out the lack of biodiversity in urban areas. To further investigate the reasons behind biodiversity loss and clarify the possible contributing factors it would be reasonable to connect land use and land coverage data to diversity indexes found in this study. The land cover information is available in Statistics Finlad website for free (Tilastokeskus).

The 16S sequencing method used in this study is commonly used for studies that focus on bacterial taxonomy and diversity. This sequencing technology gives the taxonomical data for bacteria, which is only one group of microbes. Complex microbial communities on phyllosphere and in endosphere include organisms from all domains of life, therefore it is important to understand that there can be other factors like fungi or viruses contributing to bacterial community that is cannot be recognized with 16S sequencing (Turner et al., 2013). They might reveal completely different data, although the biodiversity loss is believed to cover all kingdoms of life (Piano et al., 2020). Sequencing methods are also very sensitive for contamination. This risk of misinformation due to careless following of lab protocols or accidental contamination has to be taken into account. During this study, all of the laboratory work included negative controls which were eventually sequenced alongside actual samples. Sequences found in the negative controls were removed from the data as described in the chapter Materials and Mehtods. The results can be considered reliable in the context of laboratory work.

Better understanding of our surroundings can bring us alternative ways to balance the possible health issues arising from the lack of biodiversity around us. In this study it was shown that there is a significant difference in biodiversity between urban and rural areas, even when observed using a single indicator species. This study revealed a lower diversity of Proteobacteria wihtin urban areas compared to rural areas, which is significant, since lack of Proteobacteria diversity has been linked to immunoregulatory functions and risk of athopy (Hanski et al., 2012; Roslund et al., 2020). These results combined with scientific data linking other immune-mediated diseases with people living in highly urbanized areas (Abrahamsson et al., 2012; Hanski et al., 2012; Manichanh, 2006; Scher et al.,

2015; Rook et al.,2004) show the significance of exposing oneself into diverse nature that has often been neglected by urban planning.

The cumulating information emphasizes the importance of biodiverse environment and how the declining biodiversity affects urbanized areas. This problem is notable even in the so-called greener cities. The potential benefits from parks and other green areas are neglected because the lack of knowledge in urban planning and focus on maximizing the use of high value landscape within highly urbanized cities. However these findings can suggest that it is reasonable to think other ways of exposing urban population to rich microbiota, such as bringing environmental microbiota back to urban lifestyle by enhancing heavily built areas with biodiversity-rich materials such, as green areas, or even bringing the materials indoors (Grönroos et al., 2018; Roslund et al. ,2020; Sinkkonen et al., 2018).

6 Conclusions

This study confirmed the idea of biodiversity hypothesis (Von Hertzen et al., 2011) and how the lack of microbiological diversity is visible even in green citites where nature is present even in the close proximity of city center. The lack of biodiversity can significantly increase the risk of many immune-mediated diseases (Von Hertzen et al. 2011;Strachan, 1989; Hanski et al.2012; Strachan, 2000; Ruokolainen et al. 2016; Kondrashova 2012; Kostic et al. 2015; Von Hertzen & Haahtela, 2006), and thus present a health risk to urban dwellers. Microbial exposure from human surroundings has co-evolved hand-in-hand with humans and it seems that the proper development of the human immune system is tightly linked to that connection, (Rook, 2010; Roslund et al., 2020; Sjögren et al., 2009). Understanding of the link between microbial diversity and human health can give urban designers effective tools to prevent lack of biodiversity causing the increased risk of urban people and their pets of developing immune mediated diseases that can be prevented with increased contact with biodiversity. This can give a comprehensive idea how we should design our surroundings.

This city-ecosystem and it’s meaning towards human health needs to be studied even further. Proper understanding of the necessary exposure times and concentrations to microbial diversity can help to determine how immune mediated diseases can be prevented in the future. The exact mechanism behind microbial diversity and the normal functioning of our immune system needs also to be examined more. The understanding of environmental microbes and their importance is growing all the time, but we still have a lot to unveil. Although microbes are everywhere, but we still need to put an effort for conserving them in our everyday surroundings – even though they don’t need us. We need them.

7 Acknowledgements

I would like to thank all supporters of this master's thesis. Kiinteistö Oy Lahden Tiedepuisto granted me a 500€ scholarship during my master's thesis. Nature based-solutions research group in the Faculty of Biological and Environmental Sciences, which provided me with a topic for my thesis. Environmental Laboratory AlmaLab (University of Helsinki), which provided facilities, equipment and most importantly, personnel who gave advice and help when I needed.

My deepest gratitude goes to supervisors PhD student Laura Soininen who helped me during every step of the master's thesis process and provided endless support on thesis-related questions as well as interesting discussions including not-so-thesis-related questions including everything from houseplant care to windsocks, principle investigator Adjunct Professor Aki Sinkkonen, for supervision, constructive feedback and giving me this thesis opportunity that has led to many other amazing things, and PhD Marja Roslund for help and valuable advice during the thesis process. I would like to thank you all for your endless patience and encouragement throughout this long process.

Warm thanks to all other member of the research group: Adjunct Professor Riikka Puhakka, Dr. Mira Grönroos, PhD student Mika Saarenpää, PhD student Yan Sun and former member Dr. Anirudra Parajuli, it has been pleasure to get to know you all. I thank also the personnel in Faculty of Biological and Environmental Sciences in Lahti for your help. Special thanks for Environmental Laboratory AlmaLab laboratory analysts Kaisa Soikkeli and Pekka Keränen, who taught me all the laboratory protocols necessary for my thesis work.

Additionally, I would like to mention my family, spouse and friends who gave support and provided me with unconditional love, when I truly needed it. I cannot thank you enough.

Sources

1. Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. Austral ecology, 26(1), 32-46.
2. Anderson, M. J. (2005). Permutational multivariate analysis of variance. Department of Statistics, University of Auckland, Auckland, 26, 32-46.
3. Bar-On, Y. M., Phillips, R., & Milo, R. (2018). The biomass distribution on Earth. Proceedings of the National Academy of Sciences, 115(25), 6506-6511.
4. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), 289-300.
5. Berg, G., & Smalla, K. (2009). Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. FEMS microbiology ecology, 68(1), 1-13.
6. Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland forest communities of southern Wisconsin. Ecological monographs, 27(4), 326-349.
7. Compant S, Clement C, Sessitsch A: Plant growth-promoting bacteria in the rhizo- and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization. Soil Biol Biochem. 2010, 42: 669-678.
8. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics, 27, 2194–2200.
9. Fierer, N., Bradford, M. A., & Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. Ecology, 88(6), 1354-1364.
10. Frame, J., & Newbould, P. (1986). Agronomy of white clover. Advances in agronomy, 40, 1-88.
11. Grönroos, M., Parajuli, A., Laitinen, O. H., Roslund, M. I., Vari, H. K., Hyöty, H., ... & Sinkkonen, A. (2019). Short-term direct contact with soil and plant materials leads to an immediate increase in diversity of skin microbiota. Microbiologyopen, 8(3), e00645.
12. Gutzwiller, M. E. R., Reist, M., Peel, J. E., Seewald, W., Brunet, L. R., & Roosje, P. J. (2007). Intradermal injection of heat-killed Mycobacterium vaccae in dogs with atopic dermatitis: a multicentre pilot study. Veterinary dermatology, 18(2), 87-93.
13. Hanski, I., von Hertzen, L., Fyhrquist, N., Koskinen, K., Torppa, K., Laatikainen, T., ... & Haahtela, T. (2012). Environmental biodiversity, human microbiota, and allergy are interrelated. Proceedings of the National Academy of Sciences, 109(21), 8334-8339.

14. Hui, N., Grönroos, M., Roslund, M. I., Parajuli, A., Vari, H. K., Soininen, L., Laitinen, O. H., Sinkkonen, A., & The ADELE Research Group (2019). Diverse Environmental Microbiota as a Tool to Augment Biodiversity in Urban Landscaping Materials. Frontiers in Microbiology, 10, [536].

15. Hui, N., Parajuli, A., Puhakka, R., Grönroos, M., Roslund, M. I., Vari, H. K., ... & Sinkkonen, A. (2019). Temporal variation in indoor transfer of dirt-associated environmental bacteria in agricultural and urban areas. Environment international, 132, 105069.

16. Kallio, M. A., Tuimala, J. T., Hupponen, T., Klemelä, P., Gentile, M., Scheinin, I., ... & Korpelainen, E. I. (2011). Chipster: user-friendly analysis software for microarray and other high-throughput data. BMC genomics, 12(1), 1-14.

17. Knight, R., Vrbanac, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., ... & Melnik, A. V. (2018). Best practices for analysing microbiomes. Nature Reviews Microbiology, 16(7), 410-422.

18. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on MiSeq illumina sequencing platform. Applied and Environmental Microbiology, 79, 5112–5120.

19. Kuntaliitto, 2019. Kaupunkien ja kuntien lukumäärät ja väestötiedot [WEBPAGE]. [Referenced 31.1.2021. Available at: https://www.kuntaliitto.fi/tilastot-ja-julkaisut/kaupunkien-ja-kuntien-lukumaarat-ja-vaestotiedot]

20. Lande, R. (1996). Statistics and partitioning of species diversity, and similarity among multiple communities. Oikos, 5-13.

21. Lehtimäki, J., Karkman, A., Laatikainen, T., Paalanen, L., Von Hertzen, L., Haahtela, T., ... & Ruokolainen, L. (2017). Patterns in the skin microbiota differ in children and teenagers between rural and urban environments. Scientific reports, 7, 45651.

22. Lehtimäki, J., Sinkko, H., Hielm-Björkman, A., Salmela, E., Tiira, K., Laatikainen, T., ... & Lohi, H. (2018). Skin microbiota and allergic symptoms associate with exposure to environmental microbes. Proceedings of the National Academy of Sciences, 115(19), 4897-4902.

23. Ley, R. E., Peterson, D. A., & Gordon, J. I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell, 124(4), 837-848.

24. Lindow, S. E., & Brandl, M. T. (2003). Microbiology of the phyllosphere. Applied and environmental microbiology, 69(4), 1875-1883.

25. Luontoportti, 2021. White Clover. [Webpage]. [Referenced 6.2.2021]. Available at: https://www.luontoportti.com/suomi/en/kukkakasvit/white-clover

26. Macpherson, A. J., & Harris, N. L. (2004). Interactions between commensal intestinal bacteria and the immune system. Nature Reviews Immunology, 4(6), 478-485.

27. Manlius, Z. A. C. S. E., & LR, C. L. M. A. B. (2002). Kemeny DM Bowen G. Rook G. Walker C. Suppression of airway eosinophilia by killed Mycobacterium vaccae-induced allergenspecific regulatory T-cells. Nat Med, 8, 625-629.

28. Marquet, P. A., Fernández, M., Navarrate, S. A., & Valdovinos, C. (2004). Diversity emerging: Toward a deconstruction of biodiversity patterns. In L. Lomolino & M. Heaney (Eds.), Frontiers in biogeography: New directions in the geography of nature (pp. 191–209). Sunderland: Sinauer Associates Inc..

29. McKinney, M. L. (2002). Urbanization, Biodiversity, and Conservation The impacts of urbanization on native species are poorly studied, but educating a highly urbanized human population about these impacts can greatly improve species conservation in all ecosystems. Bioscience, 52(10), 883-890.

30. Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., & O'Hara, R.B., … Wagner, H. (2017) vegan: Community Ecology Package.

31. Oliver, A. K., Brown, S. P., Callaham, M. A. Jr, & Jumpponen, A. (2015). Polymerase matters: Non-proofreading enzymes inflate fungal community richness estimates by up to 15%. Fungal Ecology, 15, 86–89

32. Parajuli, A., Hui, N., Puhakka, R., Oikarinen, S., Grönroos, M., Selonen, V. A., ... & Nurminen, N. (2020). Yard vegetation is associated with gut microbiota composition. Science of the Total Environment, 713, 136707.

33. Parajuli, A., Grönroos, M., Siter, N., Puhakka, R., Vari, H. K., Roslund, M. I., ... & Sinkkonen, A. (2018). Urbanization reduces transfer of diverse environmental microbiota indoors. Frontiers in microbiology, 9, 84.

34. Parris, K. M. (2016). Ecology of urban environments. John Wiley & Sons.

35. Pepper, I. L., Gerba, C. P., Gentry, T. J., & Maier, R. M. (Eds.). (2011). Environmental microbiology. Academic press.

36. Piano, E., Souffreau, C., Merckx, T., Baardsen, L. F., Backeljau, T., Bonte, D., ... & Hendrickx, F. (2020). Urbanization drives cross-taxon declines in abundance and diversity at multiple spatial scales. Global change biology, 26(3), 1196-1211.

37. Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., ... & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic acids research, 41(D1), D590-D596.

38. Ravi, R. K., Walton, K., & Khosroheidari, M. (2018). MiSeq: a next generation sequencing platform for genomic analysis. In Disease Gene Identification (pp. 223-232). Humana Press, New York, NY.

39. Rook, G. A. (2009). Review series on helminths, immune modulation and the hygiene hypothesis: the broader implications of the hygiene hypothesis. Immunology, 126(1), 3-11.

40. Rook, G. A. W. (2010). 99th Dahlem conference on infection, inflammation and chronic inflammatory disorders: Darwinian medicine and the 'hygiene'or 'old friends' hypothesis. Clinical & Experimental Immunology, 160(1), 70-79.

41. Rook, G. A., Adams, V., Hunt, J., Palmer, R., Martinelli, R., & Brunet, L. R. (2004, February). Mycobacteria and other environmental organisms as immunomodulators for immunoregulatory disorders. In Springer seminars in immunopathology (Vol. 25, No. 3-4, pp. 237-255). Springer-Verlag.

42. Roslund, M. I., Grönroos, M., Rantalainen, A. L., Jumpponen, A., Romantschuk, M., Parajuli, A., ... & Sinkkonen, A. (2018). Half-lives of PAHs and temporal microbiota changes in commonly used urban landscaping materials. PeerJ, 6, e4508.

43. Roslund, M. I., Puhakka, R., Grönroos, M., Nurminen, N., Oikarinen, S., Gazali, A. M., ... & Soininen, L. (2020). Biodiversity intervention enhances immune regulation and health-associated commensal microbiota among daycare children. Science advances, 6(42), eaba2578.

44. Round, J. L., & Mazmanian, S. K. (2009). The gut microbiota shapes intestinal immune responses during health and disease. Nature reviews immunology, 9(5), 313-323.

45. Schloss, P. D., Gevers, D., & Westcott, S. L. (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PloS one, 6(12), e27310.

46. Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537-7541.

47. Sekirov, I., Russell, S. L., Antunes, L. C. M., & Finlay, B. B. (2010). Gut microbiota in health and disease. Physiological reviews, 90(3), 859-904.

48. Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). Biometrika, 52(3/4), 591-611.

49. Sinkkonen, A. T., Hyöty, H., Rajaniemi, J., Puhakka, R., Grönroos, M., Parajuli, A., Roslund, M. I., Siter, N., Nurminen, N., Oikarinen, S., Lin, J., & Laitinen, O-H. (2018). Nature-derived microbial exposure as a novel strategy to modify immune modulation among urban dwellers. Journal of Biotechnology, 280S(S12-S31), S17. https://doi.org/10.1016/j.jbiotec.2018.06.051

50. Sjögren, Y. M., Jenmalm, M. C., Böttcher, M. F., Björkstén, B., & Sverremark-Ekström, E. (2009). Altered early infant gut microbiota in children developing allergy up to 5 years of age. Clinical & Experimental Allergy, 39(4), 518-526.

51. Strachan, D. P. (1989). Hay fever, hygiene, and household size. BMJ: British Medical Journal, 299(6710), 1259.

52. Tortora, G. J., Funke, B. R., Case, C. L., & Johnson, T. R. (2004). Microbiology: an introduction (Vol. 9). San Francisco, CA: Benjamin Cummings.

53. Turner, T. R., James, E. K., & Poole, P. S. (2013). The plant microbiome. Genome biology, 14(6), 1-10.

54. University of Helsinki, 2021. Kasviatlas: Valkoapila (Trifolium repes). [Webpage]. [Referenced 6.2.2021]. Available at: http://koivu.luomus.fi/kasviatlas/maps.php?taxon=42089

55. Wilcoxon, F. (1992). Individual comparisons by ranking methods. In Breakthroughs in statistics (pp. 196-202). Springer, New York, NY.

56. Wilson, M. S., Taylor, M. D., Balic, A., Finney, C. A., Lamb, J. R., & Maizels, R. M. (2005). Suppression of allergic airway inflammation by helminth-induced regulatory T cells. The Journal of experimental medicine, 202(9), 1199-1212.

57. von Hertzen, L., & Haahtela, T. (2006). Disconnection of man and the soil: reason for the asthma and atopy epidemic?. Journal of Allergy and Clinical Immunology, 117(2), 334-344.

58. Von Hertzen, L., Hanski, I., & Haahtela, T. (2011). Natural immunity: biodiversity loss and inflammatory diseases are two global megatrends that might be related. EMBO reports, 12(11), 1089-1093.

59. Zaccone, P., Fehérvári, Z., Jones, F. M., Sidobre, S., Kronenberg, M., Dunne, D. W., & Cooke, A. (2003). Schistosoma mansoni antigens modulate the activity of the innate immune response and prevent onset of type 1 diabetes. European journal of immunology, 33(5), 1439-1449.

APPENDIXES

Appendix 1. Sampling area on map, Outer circle rural, inner circle urban